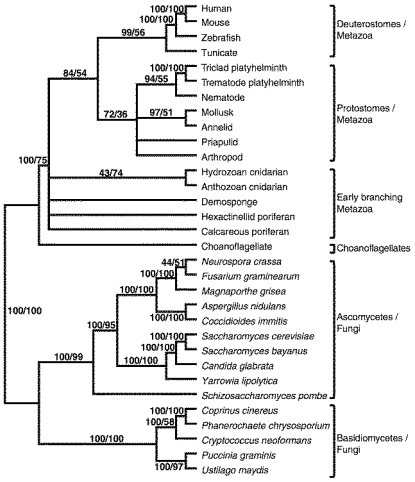
Animal evolution — A fully-resolved phylogenomic tree argues against the Cambrian explosion hypothesis

Hervé Philippe[1], Henner Brinkmann[1] and Denis Baurain[1,2]

[1]Département de biochimie, Université de Montréal, Montréal, Canada ; [2]Département des Sciences de la Vie, Université de Liège, Liège, Belgium

Abstract

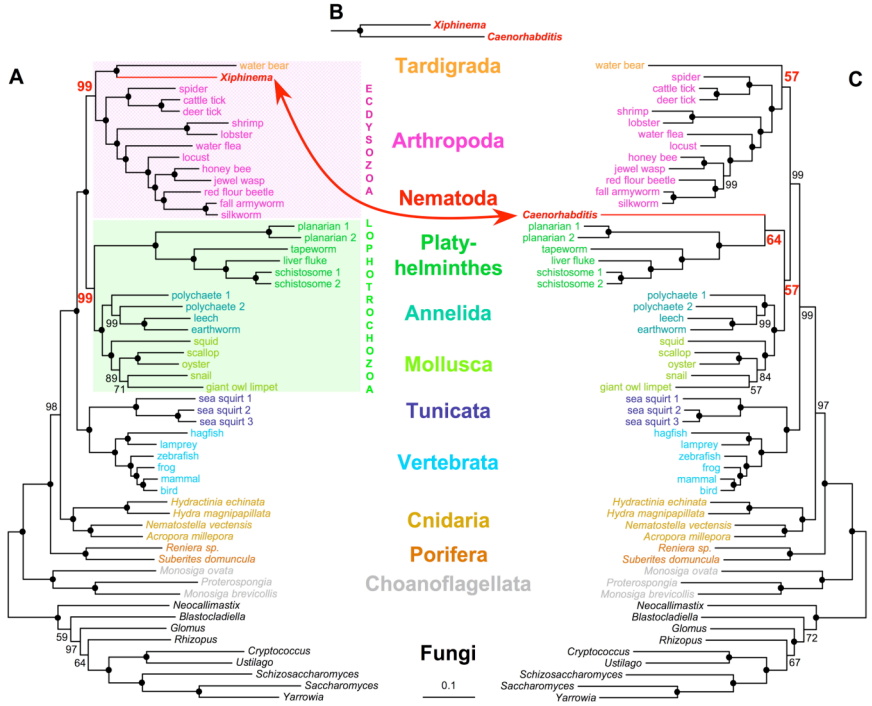
Recently, Rokas et al. *(2005) reported that the animal phylogeny could not be resolved despite the use of 50 genes from 17 animal species. Furthermore, the authors concluded that "the lack of resolution" observed in their tree constitutes a positive signature for "an extreme compression of the metazoan radiation." Hence, their work apparently supports the Cambrian explosion hypothesis. However, we suspect that the profound influence of taxon sampling on phylogenetic inference was underestimated, thus leading to erroneous conclusions about the mode and tempo of animal evolution. To substantiate this point of view, we assembled a taxon-rich phylogenomic data set. When including a slowly-evolving nematode we obtained a fully-resolved tree of animals, whereas using a fast-evolving nematode reproduced the artefacts and the lack of resolution observed by Rokas et al. The explanation for this dramatic change lies in the large amount of non-phylogenetic signal introduced by the fast-evolving nematode that annihilates the genuine phylogenetic signal. Two conclusions of general interest can be deduced from our analyses: (1) based on current data, the animal tree is fully resolved, which argues against the Cambrian explosion hypothesis; and (2) an adequate sampling (i.e. including slowly-evolving species) is crucial to reduce non-phylogenetic signals in genome-scale phylogenetic inference.*



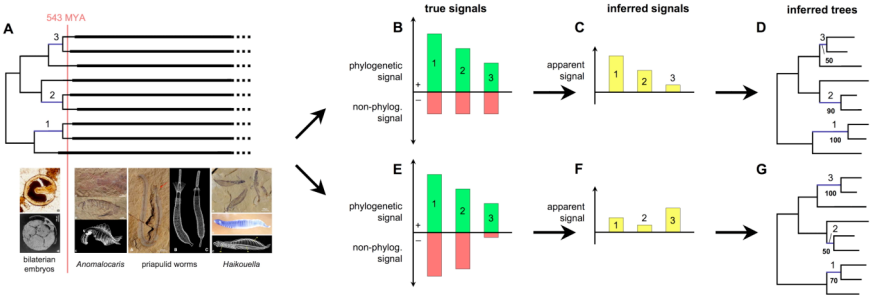
[Fig. 1] **Unresolved animal phylogeny (Rokas *et al.* 2005).** 11 out of 13 internodes in the fungal clade are significantly supported by both ML and MP analyses, whereas only 4 out of 14 internodes in the metazoan clade are significantly supported. Note the erroneous clustering of fast-evolving nematodes and plathyhelminthes, which was shown to result from a Long Branch Attraction artefact (Philippe *et al.* 2005).

Conclusions

1. **Phylogenomics allows recovery of an almost fully-resolved animal tree.** This needs further testing by including still unsampled phyla. To reduce the non-phylogenetic signals interfering with the phylogenetic signal, **slow-evolving taxa should be preferred whenever possible**.
2. Since interfering non-phylogenetic signals can be highly uneven, **internode recovery does not necessarily correlates with true internode length**.



[Fig. 2] **Phylogenomics of animals and taxon sampling.** ML trees were inferred from the concatenation of 133 genes (31,092 unambiguously aligned positions) with TreeFinder (Jobb *et al.* 2004) using a WAG+F+Γ model (100 bootstrap replicates ; closed circles denote maximum support). — **A.** When nematodes are represented by the slowly-evolving *Xiphinema*, the phylogeny is almost fully resolved and in agreement with the current consensus, recovering Ecdysozoa and Lophotrochozoa. — **B.** An excerpt of the tree inferred with both *Caenorhabditis* and *Xiphinema* shows that they form a strongly supported monophyletic group and evolve at different rates. — **C.** When nematodes are represented by the fast-evolving *Caenorhabditis*, basal nodes within protostomes are weakly supported by the bootstrap analysis. Furthermore, nematodes are artefactually attracted by the fast-evolving platyhelminthes within Lophotrochozoa (Philippe *et al.* 2005).



[Fig. 3] **The Cambrian explosion and the effects of non-phylogenetic signals on the recovery of internodes.** — **A.** The fossil record (thick lines) suggests that most animal phyla diverged at the beginning of the Cambrian. The true history is unknown. — **B.** Phylogenetic signal (PS) and non-phylogenetic signals (NPS) interact in the recovery of 3 selected nodes. The amount of PS reflects the length of the corresponding internode. NPS include, among others, rate heterogeneity and compositional biases. — **C.** Only the combined signal (CS) is available to tree reconstruction methods. Faint PS can be cancelled out by NPS (e.g. node 3). — **D.** Inferred trees depend on CS, which can affect both the length and the statistical support of internodes. — **E to G.** As above but with highly uneven NPS.

References

- G. Jobb, A. von Haeseler, K. Strimmer, *BMC Evol. Biol.* **4**, 18 (2004).
- H. Philippe, N. Lartillot, H. Brinkmann, *Mol. Biol. Evol.* **22**, 1246 (2005).
- A. Rokas, D. Kruger, S. B. Carroll, *Science* **310**, 1933 (2005).

Fossil pictures by J.-Y.Chen and J. Vannier. — D. B. is a Postdoctoral Researcher from the Fonds National de la Recherche Scientifique (FNRS, Belgium).